

---

# Orthogonal Patterns in Binary Neural Networks

---

Yoram Baram

---

(NASA-TM-100060) ORTHOGONAL PATTERNS IN  
BINARY NEURAL NETWORKS (NASA) 16 pCSCL 09B

N88-18320

Unclas  
G3/63 0125847

March 1988



National Aeronautics and  
Space Administration

---

# Orthogonal Patterns in Binary Neural Networks

---

Yoram Baram, Ames Research Center, Moffett Field, California

March 1988



National Aeronautics and  
Space Administration

**Ames Research Center**  
Moffett Field, California 94035

# ORTHOGONAL PATTERNS IN BINARY NEURAL NETWORKS

Yoram Baram\*

## Abstract

A binary neural network that stores only mutually orthogonal patterns is shown to converge, when probed by any pattern, to a pattern in the memory space--the space spanned by the stored patterns. The latter are shown to be the only members of the memory space under a certain coding condition, which allows maximal storage of  $M = (2N)^{0.5}$  patterns, where  $N$  is the number of neurons. The stored patterns are shown to have basins of attraction of radius  $N/(2M)$ , within which errors are corrected with probability 1 in a single update cycle. When the probe falls outside these regions, the error correction probability can still be increased to 1 by repeatedly running the network with the same probe.

## 1. Introduction

A mathematical model for biological neural networks, consisting of Hebb's storage mechanism [1] and McCulloch-Pitts' retrieval mechanism [2], was shown by Hopfield [3] not only to exhibit the collective behavior of the network as an associative (content-addressable) memory, but also to be technologically realizable. The model consists of  $N$  variables  $x_1, \dots, x_N$ , corresponding to the neurons of the network, each capable of having one of two values  $\pm 1$ . The state of the network is defined as the vector  $\underline{x} = (x_1, \dots, x_N)^T$  where  $(\cdot)^T$  denotes transpose. The information in  $M$  given patterns  $\underline{x}^{(1)}, \underline{x}^{(2)}, \dots, \underline{x}^{(M)}$  is stored in synaptic parameters, which are calculated according to the Hebbian rule [1]

$$T_{i,j} = \sum_{\ell=1}^M x_i^{(\ell)} x_j^{(\ell)} \quad (1.1a)$$

Information retrieval is initiated by a probe (initial state)  $\underline{x}(0)$ . Neurons are then selected at random, one at a time, and their states are updated according to the McCulloch-Pitts rule [2]

---

\*Y. Baram is a Senior Research Associate of the National Research Council at the NASA Ames Research Center, Moffett Field, CA 94035, on sabbatical leave from the Department of Electrical Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel.

$$x_i(k+1) = \text{sgn} \left\{ \sum_{j=1}^n T_{i,j} x_j(k) \right\} \quad (1.1b)$$

The network was shown by Hopfield [3] to be globally stable in the sense that, initialized by any probe, it will converge to some final state. He also observed by simulation that the stored patterns can be retrieved without severe error if  $M$  does not exceed  $0.15 N$ , for  $N = 100$ . McEllice et al. [4] and Bruce et al. [5] independently showed that the number of random patterns that can be retrieved with finite probability cannot exceed  $N/(2 \log N)$ . Hopfield further observed by simulation that "for  $N = 100$ , a pair of random memories should be separated by  $50 \pm 5$  Hamming units" [3]. When the stored patterns differ by half the bits or neuron values, they are orthogonal in the Euclidean sense. The construction of orthogonal patterns requires preprocessing or encoding of information. Neural encoding mechanisms, which involve certain notions of pattern orthogonalization, have been suggested by Kohonen [6] and by Grossberg [7]. Decoding coded patterns by neural networks has been suggested by Platt and Hopfield [8] for communication purposes and by Chiueh and Goodman [9] for pattern classification.

In this paper we first observe that when the stored patterns are mutually orthogonal, they are equilibrium states of the network (1.1). Then we propose a slight modification of the model, which allows storage only of mutually orthogonal patterns. The network state, initialized by any probe, is shown to converge to a pattern in the space spanned by the stored patterns, which we call the memory space. There can be, at most,  $N$  orthogonal patterns. It is shown that when the stored patterns satisfy a certain coding condition, they are the only members of the memory space. The maximum number of such code words is shown to be  $(2N)^{0.5}$ , which is in agreement with Hopfield's empirical observation. A particular code construction method is proposed. A network loaded with such code acts as a decoder. The stored patterns are shown to have basins of attraction of radius  $N/(2M)$ . When initialized within this range of a stored pattern, the network state converges with probability 1 to that pattern in less than a neural update cycle time. When the probe falls outside this range, the probability of retrieving the nearest stored pattern can still be increased to 1 by repeatedly running the network with the same probe.

## 2. The Memory Space

For a stored pattern to be retrievable, it is necessary that it is an equilibrium point of the network, that, once reached is never left. From (1.1) we have

$$T_{\underline{x}}(\underline{x}) = \sum_{i=1}^M \underline{x}^{(i)} [\underline{x}^{(i)}] T_{\underline{x}}(\underline{x})$$

Let the Hamming distance between  $\underline{x}^{(i)}$  and  $\underline{x}^{(l)}$  be denoted by  $d[\underline{x}^{(i)}, \underline{x}^{(l)}]$  and let  $r[\underline{x}^{(i)}, \underline{x}^{(l)}] = (1/N)d[\underline{x}^{(i)}, \underline{x}^{(l)}]$ . It can be readily verified that

$$[\underline{x}^{(i)}]^T \underline{x}^{(l)} = N - 2d[\underline{x}^{(i)}, \underline{x}^{(l)}] = N\{1 - 2r[\underline{x}^{(i)}, \underline{x}^{(l)}]\}$$

hence

$$\begin{aligned} T_{\underline{x}}^{(l)} &= \sum_{i=0}^M N\{1 - 2r[\underline{x}^{(i)}, \underline{x}^{(l)}]\} \underline{x}^{(i)} \\ &= N\underline{x}^{(l)} + \sum_{i \neq l} N\{1 - 2r[\underline{x}^{(i)}, \underline{x}^{(l)}]\} \underline{x}^{(i)} \end{aligned} \quad (2.1)$$

It can be seen that  $T_{\underline{x}}^{(l)}$  has the same sign as  $\underline{x}^{(l)}$ , unless the second term on the right-hand side of (2.1) offsets the first. Such offset cannot happen if

$$r[\underline{x}^{(i)}, \underline{x}^{(l)}] = \frac{1}{2} \quad \text{for all } i \neq l \quad (2.2)$$

This is a sufficient condition for the stored patterns to be equilibrium points of the network. It is not difficult to see that this condition is equivalent to the orthogonality condition

$$\underline{x}^{(l)T} \underline{x}^{(k)} = 0 \quad k \neq l$$

A question of interest is, how can the orthogonality condition be maintained by the neural network. Denoting by  $T(n)$  the matrix of synaptic parameters corresponding to  $n$  stored patterns, let us modify the storage rule to be

$$T(n+1) = \begin{cases} T(n) + \underline{x}^{(n+1)} \underline{x}^{(n+1)T} & \text{if } T(n) \underline{x}^{(n+1)} = 0 \\ T(n) & \text{otherwise} \end{cases} \quad (2.3a)$$

and denoting

$$z_i = \sum_{j=1}^N T_{i,j} x_j$$

let us modify the neuron update rule to be

$$x_i = \begin{cases} +1 & \text{if } z_i > 0 \\ -1 & \text{if } z_i < 0 \\ x_i & \text{if } z_i = 0 \end{cases} \quad (2.3b)$$

These slight modifications of (1.1a) and (1.1b) have the physical interpretation that the synaptic parameters remain unchanged by the probe if there is an energy release by neural firing activity. This can only occur according to (2.3b) if the probe is not orthogonal to some of the previously stored patterns. If the probe is orthogonal to all the stored patterns, that is, when it is in the null space of  $T$ , energy release by firing cannot take place and, instead, relief of potential energy is provided by a change in the synapses, meaning that the probe is stored as a new pattern. Near orthogonality may be represented by the condition  $|z_i| < \epsilon$  for a small integer  $\epsilon$ . In the rest of the paper we assume strict orthogonality of the stored patterns for mathematical simplicity. We next show that complete energy release means convergence into the space spanned by the stored patterns, which, according to the mechanism (2.3b), cannot occur if the probe is orthogonal to these patterns.

The state space of the network is the collection of all vectors of dimension  $N$ , whose components have values  $\pm 1$ . We define the network's memory space, denoted by  $X$ , as the subspace of the state space, which is spanned by the stored patterns, that is, the vectors in the state space that can be obtained as linear combinations of the stored patterns. The orthogonal projection of an arbitrary pattern  $\underline{x}$  on a stored pattern  $\underline{x}^{(l)}$  is given by

$$\hat{\underline{x}}_l = \frac{\underline{x}^T \underline{x}^{(l)}}{||\underline{x}^{(l)}||^2} \underline{x}^{(l)} = \frac{1}{N} \underline{x}^T \underline{x}^{(l)} \underline{x}^{(l)}$$

where  $||\underline{x}|| = (\underline{x}^T \underline{x})^{0.5}$  denotes the Euclidean norm of  $\underline{x}$ . It can be seen that

$$T\underline{x} = \sum_{l=1}^M \underline{x}^{(l)} \underline{x}^{(l)T} \underline{x} = N \sum_{l=1}^M \hat{\underline{x}}_l = N \hat{\underline{x}}$$

where  $\hat{\underline{x}}$  is the projection of  $\underline{x}$  on  $X$  and the last equality follows from the fact that the stored patterns  $\underline{x}^{(l)}$ ,  $l = 1, \dots, M$  are mutually orthogonal.

The energy function

$$E(\underline{x}) = -\underline{x}^T T \underline{x}$$

can be seen to have the value

$$E[\underline{x}^{(\ell)}] = - \sum_{\ell} \underline{x}^{(\ell)T} \underline{x}^{(\ell)} \underline{x}^{(\ell)T} \underline{x}^{(\ell)} = -N^2$$

for each of the stored patterns. For an arbitrary pattern  $\underline{x}$ , it has the value

$$E(\underline{x}) = -N \underline{x}^T \hat{\underline{x}}$$

By the Cauchy-Schwartz inequality

$$\underline{x}^T \hat{\underline{x}} \leq ||\underline{x}|| \cdot ||\hat{\underline{x}}||$$

Suppose that  $\underline{x}$  does not belong to  $X$ , then

$$||\hat{\underline{x}}|| < ||\underline{x}||$$

yielding, since  $||\underline{x}|| = N^{0.5}$

$$\underline{x}^T \hat{\underline{x}} < N$$

hence,

$$E(\underline{x}) > -N^2$$

It follows that for a state  $\underline{x}$  that does not belong to  $X$ , the energy is not minimal. On the other hand, if  $\underline{x}$  belongs to  $X$  (not necessarily  $\underline{x} = \underline{x}^{(\ell)}$ ), then

$$\hat{\underline{x}} = \underline{x}$$

yielding

$$E(\underline{x}) = -N \underline{x}^T \underline{x} = -N^2$$

We have shown that the minimal value of the energy function is  $-N^2$ . All points in the memory space  $X$  have this minimal energy value, while points outside the memory space have higher energy. It was shown by Hopfield [3] that the energy decreases along any path in the state space of his network. McEliece et al. [4] elaborated on Hopfield's analysis and showed that the energy can remain unchanged for only a finite number of steps. This implies that the Hopfield network converges to a point of minimal energy. Since the algorithm (2.3b) differs from the McCulloch-Pitts algorithm only by the transformation of points for which, initially,  $Tx = 0$  into new stored patterns having minimum energy, it converges, as the Hopfield model, to minimum energy points. It follows that for orthogonal stored patterns the network will converge to a point in the memory space.

### 3. Perfect Storage

We next show that under a certain condition on the stored patterns the memory space contains only these patterns. This situation may be characterized as "perfect storage." The stored patterns can then define a set of code words or, simply, a code that can be used for information representation. Suppose that the scalars  $c_1, c_2, \dots, c_M$  satisfy the equation

$$c_1 \underline{x}^{(1)} + c_2 \underline{x}^{(2)} + \dots + c_M \underline{x}^{(M)} = \underline{x} \quad (3.1)$$

where  $\underline{x}^{(1)}, \dots, \underline{x}^{(M)}$  are the stored patterns and  $\underline{x}$  is a permissible state of the network, that is, a vector whose components have values  $\pm 1$ . Then, defining the matrix

$$A = [\underline{x}^{(1)} : \underline{x}^{(2)} : \dots : \underline{x}^{(M)}]$$

and the vector

$$\underline{c} = [c_1, c_2, \dots, c_M]^T$$

we have

$$A\underline{c} = \underline{x} \quad (3.2)$$

yielding

$$\underline{c}^T A^T A \underline{c} = \underline{x}^T \underline{x} = N \quad (3.3)$$

But since

$$A^T A = NI_M$$



where  $I_M$  is the  $M \times M$  identity matrix, it follows from (3.3) that

$$\underline{c}^T \underline{c} = c_1^2 + c_2^2 + \dots + c_M^2 = 1 \quad (3.4)$$

Denoting by  $x_j^{(i)}$  the  $j$ th element of  $\underline{x}^{(i)}$ , the square of the  $j$ th row of (3.1) can be written as

$$\sum_{i,k} x_j^{(i)} x_j^{(k)} c_i c_k = 1$$

or

$$\sum_i c_i^2 + \sum_{i \neq k} x_j^{(i)} x_j^{(k)} c_i c_k = 1$$

which, by (3.4) yields

$$\sum_{i \neq k} x_j^{(i)} x_j^{(k)} c_i c_k = 0 \quad (3.5)$$

The memory space will contain only the stored patterns if the only solution to (3.5) is

$$c_i c_k = 0 \quad \text{for } i, k = 1, \dots, M, \quad i \neq k \quad (3.6)$$

which means that, at most, one of the coefficients  $c_i$  is non-zero. Defining the scalars

$$y_j^{(i,k)} = x_j^{(i)} x_j^{(k)} \quad (3.7)$$

the vectors

$$\underline{y}^{(i,k)} = [y_1^{(i,k)} \ y_2^{(i,k)} \ \dots \ y_N^{(i,k)}]^T \quad (3.8)$$

and the matrix

$$Y = [\underline{y}^{(1,2)} : \underline{y}^{(1,3)} : \dots : \underline{y}^{(1,M)} : \underline{y}^{(2,3)} : \underline{y}^{(2,4)} : \dots : \underline{y}^{(M-1,M)}] \quad (3.9)$$

and also denoting

$$\underline{c} = [c_1 c_2 \ c_1 c_3 \ \cdots \ c_1 c_M \ c_2 c_3 \ c_2 c_4 \ \cdots \ c_{M-1} c_M]^T \quad (3.10)$$

equation (3.5) can be written for all  $j, i, k$  as

$$Y \underline{c} = 0 \quad (3.11)$$

This equation has only the zero solution (3.6) if and only if  $Y$  has full column rank. The number of columns in  $Y$  is given by

$$M - 1 + M - 2 + \dots + 1 = M(M - 1)/2$$

It follows that the memory space will contain only the stored patterns if the condition

$$\text{col. rank } Y = M(M - 1)/2 \quad (3.12)$$

is satisfied. Since the number of rows of  $Y$  is  $N$ , a necessary condition for  $Y$  to have full column rank is

$$M(M - 1) \leq 2N \quad (3.13)$$

For large  $M$ , condition (3.13) may be written as

$$M \leq (2N)^{0.5} \quad (3.14)$$

We note that for  $N = 100$ , the latter condition takes the value  $M \leq 14$ , which is in agreement with the capacity bound obtained empirically by Hopfield [3]. Provided that the latter condition is satisfied, condition (3.12) is not guaranteed to hold for every choice of orthogonal patterns, as illustrated by the following examples. Consider first the orthogonal patterns formed by the columns of the matrix (where the symbol 1 has been omitted)

$$H = \begin{bmatrix} + & + & + \\ + & - & + \\ + & + & - \\ + & - & - \end{bmatrix}$$

which defines a Hadamard code  $H_{4,3}$  (see, e.g., [10], pg. 44). It can be readily verified that if the stored patterns are the first two columns of  $H$ , they are the only patterns in the memory space. Indeed, if

$$c_1 \underline{x}_1 + c_2 \underline{x}_2 = \underline{x}$$

where  $\underline{x}$  is permissible, then

$$(c_1 + c_2)^2 = 1$$

and

$$(c_1 - c_2)^2 = 1$$

yielding

$$c_1 c_2 = 0$$

Hence, either  $c_1 = 0$  or  $c_2 = 0$  and since  $c_1^2 + c_2^2 = 1$ , the assertion follows. Similarly, if all three columns of  $H$  are stored, they are the only patterns in the memory space, as the product matrix  $Y$ , obtained as

$$Y = \begin{bmatrix} + & + & + \\ - & + & - \\ + & - & - \\ - & - & + \end{bmatrix}$$

can be seen to have mutually orthogonal columns.

Next, suppose that the stored patterns are the columns of the matrix

$$H = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \\ + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}$$

Which are the first four columns of the Hadamard matrix obtained by the Sylvester matrix construction (see [10], p. 45). It can be seen that the product matrix

$$Y = \begin{bmatrix} + & + & + & + & + & + \\ - & + & - & - & + & - \\ + & - & - & - & - & + \\ - & - & + & + & - & - \\ + & + & + & + & + & + \\ - & + & - & - & + & - \\ + & - & - & - & - & + \\ - & - & + & + & - & - \end{bmatrix}$$

does not have full column rank (the third and the fourth columns are the same and so are the second and the fifth and the first and the sixth). On the other hand, replacing the last columns in  $H$  by another, as in

$$H = \begin{bmatrix} + & + & + & + \\ + & - & + & + \\ + & + & - & + \\ + & - & - & + \\ + & + & + & - \\ + & - & + & - \\ + & + & - & - \\ + & - & - & - \end{bmatrix} \quad (3.15)$$

yields the product matrix

$$Y = \begin{bmatrix} + & + & + & + & + & + \\ - & + & + & - & - & + \\ + & - & + & - & + & - \\ - & - & + & + & - & - \\ + & + & - & + & - & - \\ - & + & - & - & + & - \\ + & - & - & - & - & + \\ - & - & - & + & + & + \end{bmatrix}$$

which can be seen to have mutually orthogonal columns.

The above example shows that the satisfaction of condition (3.12) is not guaranteed by orthogonality of the code words but, rather, depends on the particular

choice of the code. The question arises whether there is a systematic code construction that guarantees satisfaction of the condition. We next suggest such a construction. Let  $N = 2^k$  for some positive integer  $k$  and suppose, without loss of generality, that the first code word is  $(+ + \dots +)^T$ . Divide the first word into two equal sections and change the signs of the bits in one section. Divide each of these sections again into two equal sections and change the signs in one. This operation can be repeated only  $k$  times and results in  $k + 1$  mutually orthogonal code words. Since the bit-wise products of the resulting code words maintain the original division modulo sign change, the resulting product words are mutually orthogonal. Hence, the desired condition is satisfied. Such code construction for  $N = 8$  yields

$$H = \begin{bmatrix} + & + & + & + \\ + & + & + & - \\ + & + & - & + \\ + & + & - & - \\ + & - & + & + \\ + & - & + & - \\ + & - & - & + \\ + & - & - & - \end{bmatrix}$$

which can be seen to include the same code words as (3.15). The latter was shown above to satisfy the condition.

#### 4. Error Correction

We have seen that under condition (3.12) the memory space contains only the stored patterns. Consequently, the network, probed by any pattern, will converge to one of the stored patterns. If the latter are viewed as code words, the probe may be viewed as such a word corrupted by noise. Convergence to the code word closest to the probe in Hamming distance then means error correction. We next show that the network has the capability of correcting a certain number of bits, with probability 1. This number defines the "basins of attraction" of the stored patterns. Suppose that  $\underline{x}^{(l)}$  is the stored pattern closest to the probe  $\underline{x}$ . We have

$$\begin{aligned}
T_{\underline{x}} &= \sum_{j=1}^M \{N - 2d[\underline{x}^{(j)}, \underline{x}]\} \underline{x}^{(j)} \\
&= \{N - 2d[\underline{x}^{(\ell)}, \underline{x}]\} \underline{x}^{(\ell)} + \sum_{j \neq \ell} \{N - 2d[\underline{x}^{(j)}, \underline{x}]\} \underline{x}^{(j)} \quad (4.1)
\end{aligned}$$

Let us use the abbreviated notation  $d = d[\underline{x}^{(\ell)}, \underline{x}]$ . Suppose that the  $i$ 'th neuron is to be updated. The distance  $d$  will decrease or remain the same if and only if the sign of  $[T_{\underline{x}}]_i$  is the same as that of  $[\underline{x}^{(\ell)}]_i$ . In the worst case all  $[\underline{x}^{(j)}]_i$ ,  $j \neq \ell$  have the same sign, opposite to that of  $[\underline{x}^{(\ell)}]_i$ , and

$$d[\underline{x}^{(j)}, \underline{x}] = \frac{1}{2} N - d \quad \text{for all } j \neq \ell$$

which yields the maximum offset of the first term on the right-hand side of (4.1) by the second. In this situation, the sign of  $[T_{\underline{x}}]_i$  will be the same as that of  $[\underline{x}^{(\ell)}]_i$  if and only if

$$(N - 2d) > (M - 1)[N - 2(\frac{1}{2} N - d)]$$

or,

$$d < \frac{N}{2M} \quad (4.2)$$

When a code word is corrupted so that, at most,  $N/(2M)$  of its bits are erroneous, it will be corrected with probability 1 by the network in a single neural update cycle. The neighborhood within a distance  $N/(2M)$  about a stored pattern is its "basin of attraction." When the probe falls outside this range, the network's state may still converge with high probability to the closest stored pattern, depending on its distance from the probe.

Suppose that the network is rerun repeatedly with the same probe. The final state of the first run is registered. The final state of each consecutive run replaces that of the previous one if it is closer to the probe and is discarded otherwise. Since the neuron selections for update are mutually independent, so are the resulting state trajectories in the different network runs. Each run terminates in one of the stored patterns. Since the probability of converging to the pattern closest to the probe is at least as large as that of converging to any other stored pattern, it is finite. It follows that the probability of recovering the stored pattern closest to the probe increases to 1 as the number of runs is increased to infinity.

## 5. Conclusion

The storage of mutually orthogonal patterns in a binary neural network guarantees convergence of the network state, initialized by any pattern, to a pattern in the memory space. Under a certain coding condition, the memory space contains only the stored patterns. The state converges to the nearest stored pattern with probability 1 when it is initialized within the latter's basin of attraction. Otherwise, the probability of error correction can be increased asymptotically to 1 by repeatedly running the network with the same probe.

## References

1. D.O. Hebb, The Organization of Behaviour, Wiley, New York, 1949.
2. W.S. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Imminent in Nervous Activity," Bull. Math. Biophys., Vol. 5, pp. 115-133, 1943.
3. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proc. National Academy of Sciences, Vol. 79, pp. 2554-2558, April 1982.
4. R.J. McEliece, E.C. Posner, E.R. Rodemich and S.S. Venkatesh, "The Capacity of the Hopfield Associative Memory," IEEE Trans. on Information Theory, Vol. IT-33, No. 4, July 1987.
5. A.D. Bruce, E.J. Gardner and D.J. Wallace, "Dynamics and Statistical Mechanics of the Hopfield Model," J. Phys. A: Math. Gen., Vol. 20, pp. 2909-2934, 1987.
6. T. Kohonen, Associative Memory - A System Theoretic Approach, Springer Verlag, New York, 1977.
7. S. Grossberg, Studies of Mind and Brain, D. Reidel, Boston, 1982.
8. J.C. Platt and J.J. Hopfield, "Analog Decoding Using Neural Networks," Neural Networks for Computing Conference, Snowbird, Utah, 1986.
9. T-D. Chiueh and R. Goodman, "A Neural Network Classifier Based on Coding Theory," IEEE Conference on Neural Information Processing Systems, Denver, 1987.
10. F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes, North Holland, New York, 1977.

# Report Documentation Page

1. Report No.  NASA TM-100060		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle  Orthogonal Patterns in Binary Neural Networks				5. Report Date March 1988	
				6. Performing Organization Code	
7. Author(s)  Yoram Baram				8. Performing Organization Report No.  A-88068	
				10. Work Unit No.  505-66-11	
9. Performing Organization Name and Address  Ames Research Center Moffett Field, CA 94035				11. Contract or Grant No.	
				13. Type of Report and Period Covered  Technical Memorandum	
12. Sponsoring Agency Name and Address  National Aeronautics and Space Administration Washington, DC 20546-0001				14. Sponsoring Agency Code	
15. Supplementary Notes  Point of Contact: Heinz Erzberger, Ames Research Center, MS 210-9, Moffett Field, CA 94035, (415) 694-5425 or FTS 464-5425					
16. Abstract  A binary neural network that stores only mutually orthogonal patterns is shown to converge, when probed by any pattern, to a pattern in the memory space--the space spanned by the stored patterns. The latter are shown to be the only members of the memory space under a certain coding condition, which allows maximal storage of $M = (2N)^{0.5}$ patterns, where N is the number of neurons. The stored patterns are shown to have basins of attraction of radius $N/(2M)$ , within which errors are corrected with probability 1 in a single update cycle. When the probe falls outside these regions, the error correction probability can still be increased to 1 by repeatedly running the network with the same probe.					
17. Key Words (Suggested by Author(s))  Neural networks Coding			18. Distribution Statement  Unclassified - Unlimited   Subject Category - 63		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of pages 14	
				22. Price A02	